Routledge
Taylor & Francis Group

Bilingual Research Journal

The Journal of the National Association for Bilingual Education

ISSN: 1523-5882 (Print) 1523-5890 (Online) Journal homepage: http://www.tandfonline.com/loi/ubrj20

Exploring the cross-linguistic transfer of reading skills in Spanish to English in the context of a computer adaptive reading intervention

Doris Luft Baker, Deni Lee Basaraba, Keith Smolkowski, Jillian Conry, Jarkko Hautala, Ulla Richardson, Sherryl English & Ron Cole

To cite this article: Doris Luft Baker, Deni Lee Basaraba, Keith Smolkowski, Jillian Conry, Jarkko Hautala, Ulla Richardson, Sherryl English & Ron Cole (2017): Exploring the cross-linguistic transfer of reading skills in Spanish to English in the context of a computer adaptive reading intervention, Bilingual Research Journal, DOI: 10.1080/15235882.2017.1309719

To link to this article: http://dx.doi.org/10.1080/15235882.2017.1309719

BILINGUAL RESEARCH JOURNAL
http://dx.doi.org/10.1080/15235882.2017.1309719

Routledge
Taylor & Francis Group

Exploring the cross-linguistic transfer of reading skills in Spanish to English in the context of a computer adaptive reading intervention

Doris Luft Baker[a], Deni Lee Basaraba[b], Keith Smolkowski[c], Jillian Conry[a], Jarkko Hautala[d], Ulla Richardson[d], Sherryl English[a], and Ron Cole[e]

[a]Southern Methodist University; [b]Bethel School District; [c]Oregon Research Institute; [d]University of Jyväskylä; [e]Boulder Learning Inc.

ABSTRACT

We explore the potential of a computer-adaptive decoding game in Spanish to increase the decoding skills and oral reading fluency in Spanish and English of bilingual students. Participants were 78 first-grade Spanish-speaking students attending bilingual programs in five classrooms in Texas. Classrooms were randomly assigned to the treatment (i.e., where students played Graphogame Spanish) for 16 weeks for ten minutes per day (n = 3) versus business as usual instruction (n = 2). Results indicate that students at some risk on Spanish pseudoword reading appeared to benefit the most from playing the game. Analysis of gains suggests a potentially small, but meaningful educational effect of the game on Spanish oral reading fluency and English pseudoword reading when taking Spanish decoding skills at pretest into account. Students indicated that they enjoyed playing the game, and that the game helped them improve their reading skills. Teachers perceived the game as an engaging tool for students to use during small-group instruction or during independent time in a Response-to-Intervention approach. We discuss our mixed results in the context of using computer-adaptive games to improve the academic outcomes of bilingual students.

Introduction

Acquiring word automaticity, the ability to read words quickly and effortlessly, is a prerequisite to learning to read fluently and with comprehension in alphabetic languages such as Spanish, Finnish, and English (Ehri, 2005; Kyle, Kujala, Richardson, Lyytinen, & Goswami, 2013; LaBerge & Samuels, 1974; Perfetti, 1985). According to Gough and Tunmer's (1986) simple view of reading, the ability to read words, and understand what one reads, forms the basis of reading. Key to word reading automaticity is mastering the alphabetic principle, which is composed of two essential skills: (a) phonological decoding, or the understanding that printed letters in words systematically represent the sounds of spoken language; and (b) phonological recoding, or the ability to blend sounds

CONTACT Doris Luft Baker dluftdebaker@smu.edu Department of Teaching and Learning, Simmons School of Education & Human Development, Southern Methodist University, Suite 370, Dallas, TX 75275.

Doris Luft Baker is affiliated with Teaching and Learning, Southern Methodist University, Dallas, TX.
Deni Lee Basaraba is affiliated with the Bethel School District, Eugene, Oregon.
Keith Smolkowski is affiliated with the Oregon Research Institute, Eugene, OR.
Jillian Conry is affiliated with Teaching and Learning, Southern Methodist University, Dallas, TX.
Jarkko Hautala and Ulla Richardson are affiliated with the Agora Center & Department of Psychology, University of Jyväskylä, Jyväskylä, Finland.
Sherryl English is affiliated with Teaching and Learning, Southern Methodist University, Dallas, TX.
Ron Cole is affiliated with Boulder Learning Inc., Boulder, CO.

together to read words (Adams, 1990; Ehri, 2005). Although both of these skills are fundamental for learning to read in English, their relative importance and contribution may be different in alphabetic languages with different orthographic transparencies (Crespo, 2014; Kyle et al., 2013; Seymour, 2005).

Spanish, for example, is considered to have a relatively transparent orthography in which there is close to a one-to-one correspondence between phonemes and the graphemes used to represent them. There are few exceptions, such as the vowel *u* that can be pronounced as /u/ as in *uno* but that is also silent between the letters *q* and *e* or *i* as in *queso* ("cheese") or *quiso* ("wanted") (Davies, Cuetos, & Glez-Seijas, 2007). English, in contrast, has been described as an opaque or even "irregular" language (Foorman, Breier, & Fletcher, 2003) because it requires students to use the 26 letters of the alphabet to represent the approximately 40 phonemes of the English language (Lyon, 2009). Moreover, English uses only five vowel letters to represent approximately 12–15 vowel sounds (Frost, 2005). These sounds are pronounced in many different ways depending on the vowel–vowel or consonant–vowel combinations (e.g., the vowel *o* has seven different sounds and 13 spelling forms as in *load, hold, boil, toy, boot, short, cloud, own, not, ocean, robe, toe, owl, soup*; Genessee, Geva, Dressler, & Kamil, 2006).

Regardless of these orthographic differences, acquiring word reading automaticity also requires the coordination of a series of bottom-up skills such as phonemic awareness, letter-sound correspondence knowledge, and word recognition that, for some readers, takes years to master (Kame'enui & Simmons, 2001). Ehri (2005) suggests that beginning readers progress through a series of phases to create a mental representation of words that allows them to then retrieve these words automatically from memory.

Recent evidence also indicates that mastering the alphabetic principle in Spanish or English not only provides readers with the ability to read words automatically and with fluency when reading connected text but also has a significant, direct relation with reading comprehension, not only within these languages but also *across* these languages (Baker, Park, & Baker, 2012; Baker, Stoolmiller, Good, & Baker, 2011; Clemens, Shapiro, Wu, Taylor, & Caskie, 2014). Moreover, gains in word reading automaticity account for a significant percentage of the variance in reading comprehension scores measured by standardized tests in both languages (Baker, Park, & Baker, 2010; Fien et al., 2010). For example, Baker et al. (2010) examined the relation between initial status and growth on measures of alphabetic understanding and reading comprehension in Spanish with 168 students in Grades K and 1 in the United States. Results indicated that initial status and growth in decoding pseudowords were significant direct predictors of oral reading fluency (ORF) and reading comprehension in Spanish at the end of first grade. In particular, findings suggested that a 10-point higher Spanish pseudoword reading score at the end of kindergarten predicted a 5.5-higher Spanish reading comprehension score at the end of first grade. The model (i.e., initial status and growth in pseudoword reading) explained approximately 53% of the observed variance in first-grade reading comprehension scores in Spanish. This study, however, did not examine whether initial status and growth in pseudoword reading depended on individual student differences in reading skills and did not examine whether Spanish decoding skills had an effect on student English decoding skills.

In a similar study with English-speaking students, Fien et al. (2010) examined the relation between initial status and gains in pseudoword reading and oral reading fluency on reading comprehension at the end of first grade. Findings indicated that initial status explained about 37% of the student-level variance in English reading comprehension at the end of first grade and that gains on pseudoword reading throughout the year explained an additional 9%. Moreover, gains on pseudoword reading were more predictive of later reading outcomes for struggling students than for students who were already reading fluently. The findings from these two studies illustrate the importance of supporting the development of alphabetic understanding skills as early as kindergarten and certainly in first and second grades because of its direct connection not only with oral reading fluency but also with more complex reading comprehension processes in Spanish and English.

The purpose of this study is to explore the effects of a Computer Adaptive Reading Intervention, or CARI, called GraphoGame (Lyytinen, Ronimus, Alanko, Poikkeus, & Taanila, 2007) to support the understanding of the alphabetic principle in Spanish and English for first-grade bilingual students attending one-way dual language programs (i.e., where the majority of the students in the class are Spanish speakers learning to read in Spanish in kindergarten and first grade) in Texas. We hypothesize that if CARIs, in general, can help students master the alphabetic principle, and potentially transfer these skills to their second language through engaging practice opportunities, then teachers can devote more of their instructional time to vocabulary and reading comprehension activities that are more complex and, at least up until now, have been challenging to deliver via a computer game. In the next section we discuss the research on cross-linguistic transfer, particularly between alphabetic languages such as Spanish and English.

Cross-linguistic transfer

The few studies that have examined the effects of CARIs on bilingual student reading outcomes suggest that there is a need to explore the use of CARIs to support the reading development of bilingual students not only within languages but also across languages as an effect of cross-linguistic transfer (Baker, Basaraba, & Polanco, 2016). This examination is particularly important considering the proliferation of bilingual programs in the United States (Maxwell, 2012) and the substantial evidence suggesting that bilingual programs are effective ways to support the development of English reading skills for bilingual students whose native language is not English (Baker et al., 2016). *Cross-linguistic transfer* refers to a student's ability to transfer what they know in one language, (e.g., Spanish) to another language (e.g., English). This type of transfer is based on two predominant theoretical frameworks: linguistic interdependence (Cummins, 1979) and contrastive analysis (Lado, 1957).

Linguistic interdependence posits that the process of learning a second language (L2) is connected to the underlying proficiency in the first language (L1) and that all language acquisition occurs through a system that extends beyond cognitive skills. In other words, cross-linguistic transfer can occur only if students' language and reading proficiency in their L1 is strong and the student has had some exposure to L2. On the other hand, the contrastive analysis hypothesis posits that learners compare and contrast elements of L1 with those of L2 as they acquire the second language. Studies based on these theoretical frameworks have produced mixed results, but both theories are supported by empirical evidence.

For example, several studies have indicated that L1 instruction can impact a variety of L2 language and literacy abilities such as spelling (San Francisco, Carlo, August, & Snow, 2006), phoneme segmentation (Gerber et al., 2004), reading comprehension (Nakamoto, Lindsey, & Manis, 2012), and writing fluency (Bae, 2007). Evidence also indicates that both L1 and L2 knowledge are independent predictors of L2 phonemic awareness (Verhoeven, 2011) and that cross-linguistic transfer impacts phonological awareness even when the languages use different writing systems (e.g., one alphabetic and one logographic; Luo, Chen, & Geva, 2014). Moreover, cross-linguistic transfer can also impact other components of reading across languages. For example, Nakamoto et al. (2012) found that Spanish decoding skills in first grade mediated the relation between Spanish decoding skills in kindergarten and English reading fluency in third grade. In another study, Baker, Park, and Baker (2013) found that Spanish word reading and oral reading fluency in first grade had a significant effect on reading comprehension in English at the end of second grade.

In addition, a recent meta-analysis by Melby-Lervåg and Lervåg (2011) suggests that cross-linguistic correlations of decoding and phonological awareness appear to be stronger (r = .54 and r = .44) than correlations in the oral language domain (r = .16). In summary, a large body of evidence suggests an interconnectedness among languages. However, few studies have experimentally investigated cross-linguistic transfer through the use of CARIs. Next we review the evidence of using CARIs in educational settings.

Computer adaptive reading interventions (CARIs)

CARIs offer several advantages over traditional, teacher-directed instruction, including the ability to (a) adapt to the specific needs of diverse students, (b) present content in a dynamic and engaging format with graphics that are not available in traditional materials, (c) allow students to control the pace of their own learning, and (d) provide immediate individualized feedback regarding responses (Potocki, Ecalle, & Magnan, 2013). Moreover, due to increasing access to technology within schools, the possibility of implementing CARIs is no longer restricted to schools serving students from higher socioeconomic backgrounds (State Educational Technology Directors Association [SETDA], 2010).

Thus, the availability of technological resources (e.g., desktop computers, laptops, tablet computers), coupled with devices that allow Internet access at minimal cost to users, means that schools have increasingly more opportunities to use technology to support all learners than in previous years. Nemeth and Simon (2013) suggest that access to the Internet provides bilingual students with translation tools, online dictionaries, and educational game applications that empower them to take learning into their own hands and communicate more efficiently with their teachers. Moreover, since 2000, several studies and literature reviews have examined the efficacy of CARIs for supporting the reading achievement of young learners. For example, results of a meta-analysis designed to examine the effectiveness of CARIs in improving the reading skills of students with reading disabilities suggested that CARIs that incorporated systematic instructional procedures (e.g., immediate feedback, instructive and consistent correction procedures, multiple opportunities to practice, individualized pacing) helped overcome the "Matthew Effect" (Stanovich, 1986) of early literacy development (i.e., where children with higher literacy skills improve more and faster than children with lower literacy skills; Hall, Hughes, & Filbert, 2000).

Two relevant studies (Lan, Sung, & Chang, 2009; Macaruso & Rodman, 2011) have investigated the effectiveness of phonics-based CARIs with bilingual students. Lan et al. (2009) examined the effectiveness of a computer-assisted reciprocal early English reading system (CAREER) for supporting the English reading development of 52 students in four Grade 4 classrooms randomly assigned to either the treatment (CAREER) or control condition. Over 10 weeks students in both conditions received approximately 160 minutes of English reading instruction focused on sight words, decodable words, and text reading. Results indicated that students in the treatment condition had higher scores on Retell Fluency compared to students in the control condition, although ORF scores were similar in both conditions after controlling for student English proficiency scores.

Similarly, Macaruso and Rodman (2011) investigated the effectiveness of the *Lexia Reading* software for supporting the early literacy development of Spanish-speaking kindergarten students in Texas. Findings indicated that students who participated in a minimum of 45 sessions of the CARI made significantly greater gains on the Group Reading Assessment and Diagnostic Evaluation (GRADE, Williams, Cassidy, & Samuels, 2001) total test score, the word reading subtest, and the phonological awareness subtest than their peers in the control condition who received instruction in math and other activities. Recent studies have also indicated that CARIs and virtual gaming can play an important role in language learning because of their competitive nature, tendency to follow storylines, appealing interactivity, and production of circumstances that elicit communication. Moreover, the colorful, audiovisual nature of digital games attracts students' attention and seems to foster sustained engagement (Rusu-Bodea, 2016). Few studies, however, have experimentally tested these hypotheses.

A review by Cheung and Slavin (2013) of 20 studies published between 1980 and 2010 that included approximately 7,000 students in Grades 1–6 suggested that CARIs may be more effective with struggling readers in Grades 1–3 than with older struggling readers, with an overall effect size of +0.36 for Grades 1–3 and an effect size near 0 (+0.07) for Grades 4–6. In summary, results of empirical studies on the use of CARIs to improve reading and language outcomes indicate that CARIs have the potential to engage students quickly in reading and language activities that are conducive to learning. However, results of Cheung and Slavin's meta-analysis and our own review of

the literature suggest that it is still not clear how, and under what circumstances, CARIs improve student reading performance, particularly for students attending bilingual programs.

Evidence supporting the effectiveness of GraphoGame

Given that the purpose of this study is to explore the use of a CARI in the context of students learning to read in Spanish and in English, next we examine the efficacy of GraphoGame in different settings and with students outside the United States. GraphoGame is an adaptive online decoding program developed by the University of Jyväskylä (Finland) and the Niilo Mäki Foundation (Jyväskylä, Finland). The game was originally developed in Finnish as a promising tool to assess and intervene with beginning readers who were struggling with learning letter-sound correspondences in Finnish (see Lyytinen et al., 2007; Saine, Lerkkanen, Ahonen, Tolvanen, & Lyytinen, 2011). The game is called Ekapeli in Finnish and is available online free of charge to all learners (see http://info.graphogame.com/). Currently, more than 250,000 students have used the Finnish version of the game, and thousands of students have played the game in more than 20 different languages.

To examine the efficacy of GraphoGame, Saine et al. (2011) conducted a randomized control trial with 166 seven-year-old students identified as having reading difficulties. Students who were below the 30th percentile on a screening reading measure were randomly assigned to either a remedial reading instructional group (RRI) or to the GraphoGame group for 28 weeks. Results indicated that students who played GraphoGame had significantly higher scores and made significantly greater gains than their peers in the RRI group on all reading measures (*ES* = + 2.08, +1.01, and +1.68 for letter knowledge, reading accuracy, and spelling respectively). More importantly, these gains were sustained 12 months and even 16 months after the intervention. In addition, students who played GraphoGame were able to catch up to the mainstream reading group, closing the gap between struggling readers and proficient readers. Other studies in different languages such as Polish, German, and several different African languages have shown the promise of GraphoGame to improve children's decoding skills (see Jere-Folotiya et al., 2014; Kamykowska, Haiman, Latvala, Richardson, & Lyytinen, 2013; Ojanen et al., 2015; Ronimus & Richardson, 2014). However, some of these studies have included a small sample size, and students were not randomly assigned to treatment and control groups (Kyle et al., 2013), or they were correlational (Ronimus & Richardson, 2014). Despite these limitations, the game has been well received by teachers and students across the globe (Ojanen et al., 2015).

Thus, the purpose of this study is to explore the effects of GraphoGame in Spanish (GG-Spanish) by conducting a small-scale descriptive evaluation, supplemented by statistical analysis of the impact of the game, in the context of bilingual programs in the United States. Specifically, we attempted to answer the following research questions:

RQ1: Do first-grade bilingual students in the United States who played GraphoGame in Spanish (GG-Spanish) perform better on measures of decoding and oral reading fluency at posttest in Spanish compared to first-grade bilingual students who did not play GG-Spanish? Is there a difference in outcomes based on student risk status at pretest?

RQ2: To what extent does playing GG-Spanish support the cross-linguistic transfer of decoding and ORF skills in Spanish to English? Does the support of cross-linguistic transfer depend on student risk status at pretest?

RQ3: Did students find GG-Spanish easy to use and engaging?

RQ4: Did teachers find GG-Spanish easy to incorporate into their regular classroom routines?

To respond to RQ1 and RQ2, we used an experimental design because we wanted to obtain initial evidence of the effect of GG-Spanish in the United States. Given that Graphogame was developed to support mostly students at risk for reading difficulties, we also decided to look at differences in outcomes based on student reading risk status, which was determined by the Indicadores Dinámicos del Exito en la Lectura (IDEL) 2007 benchmark goals (Baker, Cummings, Good, & Smolkowski, 2007) and by the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) 2012 benchmark goals (University of Oregon Center on Teaching & Learning, 2012) in the middle and at the end of first grade. RQ3 and RQ4 were answered using an observation tool, surveys, and interviews, and they were designed to determine whether GG-Spanish was a feasible activity in the context of U.S. bilingual classrooms.

Method

Sample and research design

Prior to beginning the intervention, we looked for schools with bilingual programs in a large metropolitan area in Texas that would be interested in collaborating with us on this project. Our convenience sample consisted of three schools located in a high-poverty neighborhood, as indicated in Table 1, and five first-grade teachers who provided Spanish reading instruction in a one-way dual language model. All teachers were native Spanish speakers certified in Elementary Education and Bilingual Education. They had an average of eight years of overall teaching experience, with an average of 4.5 years of experience teaching first grade. Eighty bilingual students (53% male) in the five classrooms returned their parental consent form to participate in the study. All students were Hispanic, and 76.9% were identified as Limited English Proficient.

We used an experimental design in which classrooms were randomly assigned within schools to playing GG-Spanish for 16 weeks (n = 3) versus "business as usual" instruction (i.e., without playing GG-Spanish; n = 2). We randomly assigned classrooms instead of students to receive GG-Spanish versus business as usual because we were interested in manipulating variables (i.e., the use of a game and the language transfer) instead of only looking at correlations between pretest and posttest without a control group. In addition, using the classroom as the unit of analysis reduced the potential frustration of students in the control group for feeling left out from playing a game that has been found to be very engaging and motivating by children across the world (see Ojanen et al., 2015; Ronimus & Richardson, 2014). Forty-seven students were in the treatment classrooms, and 33 were in the control classrooms. Two students (one in the treatment and one in the control group) did not have complete data due to absences. Therefore, our analytic sample consisted of 78 students.

Description of GG-Spanish

The Spanish version of GraphoGame was originally developed for an intervention study in Chile and is currently the only version of the game in Spanish. The game prompts the player to listen to spoken sounds (corresponding to one or more phonemes) and then use a mouse to click on the corresponding letter or set of letters. After an incorrect selection, the game immediately shows the correct grapheme for the presented sound. Based on student performance, the game provides learning

Table 1. Demographic information by school (percentages of students).

	White	Black	Hispanic	Other	Receiving FRL	LEP	SPED
School A	2.3	31.8	65.5	0.4	100.0	52.5	5.4
School B	1.1	9.3	88.2	1.4	97.5	41.0	3.8
School C	3.9	12.9	81.5	1.8	92.0	42.7	8.5

Note. FRL = Free and Reduced-Price Lunch, LEP = Limited English Proficiency, SPED = Students receiving special education services.

material in the subsequent trials or levels that allow the player to select the correct response an average of 80% of the time. In addition, similar game levels are presented in several different graphical settings to provide students with multiple practice opportunities of the same task while simultaneously maintaining their engagement. Each game level is relatively short, lasting, on average, from one to three minutes.

In this study, bilingual students played the game in Spanish during language arts in either the school computer lab or the classroom. When the game was played in the computer lab, the teacher circulated around the room to make sure all students were logged in correctly and that they were using the headphones appropriately. After approximately 10 minutes, a timer beeped, and students returned to the classroom. When students played GG-Spanish in the classroom, the teacher assigned four or five students to a computer. Students logged in and started playing the game, and each time the timer beeped, a new group of students got on the computers and played the game. Meanwhile, the teacher was providing small-group instruction to another four to five students, and the rest of the students were engaged in independent activities such as reading, drawing, or playing word games.

Professional development

Teachers in the treatment group received a two-hour training on how to set up the program to help students who would have a difficult time logging in with their ID number and to ensure that all students in the class had an approximately equal amount of time to play the game. During the training, teachers were also able to log on to the game Web site, play the game, and participate in the different levels that students would progress through as they played. Teachers asked clarifying questions and received feedback and assistance from the research staff. Additional technical support, if needed, was provided throughout the study.

Measures

All students were assessed with measures of pseudoword reading and oral reading fluency in Spanish and in English before and after the intervention. In addition, students in the treatment group completed a survey to elicit information about their perceptions of the usability of, and level of engagement with, GG-Spanish. Teachers completed a survey to evaluate feasibility and report their impressions of level of student engagement. We also conducted observations in the treatment and control classrooms. Next we describe each measure in detail.

IDEL fluidez en las Palabras sin Sentido (IDEL FPS; Plasencia- Peinado, Baker, Good, & Peinado, 2006)

FPS is a standardized, individually administered test of pseudoword reading in Spanish. It is a subtest of the Indicadores Dinámicos del Exito en la Lectura (Baker, Good, Knutson, & Watson, 2006). Successful performance on FPS indicates knowledge of (a) letter-sound correspondences, and (b) how to blend letter-sounds into whole units (i.e., pseudowords). Three-week, alternate-form reliability of FPS in the middle of first grade was .76 (Baker, Good, Peyton, & Watson, 2004). The concurrent validity of FPS with the Woodcock-Muñoz Pruebas de Aprovechamiento subtest of Análisis de Palabras was .72 at the end of first grade (Watson, 2004). In this study, we administered alternate forms of FPS at pretest and posttest as a proximal outcome.

IDEL Fluidez en la Lectura Oral (IDEL FLO; Baker, Good, Mross, et al., 2006)

FLO is a standardized, timed, individually administered test of accuracy and fluency in reading connected text in Spanish (Baker, Good, Knutson, & Watson, 2006). Passages were written taking into account sentence length, number of high-frequency words, and number of letters and syllables in words. Students read each passage for 1 minute. Words omitted or substituted, and hesitations of more than 3 seconds, are scored as errors. Alternate-form reliability of different reading passages

from the same level of difficulty ranged from .88 to .94 (Baker et al., 2004). Criterion-related validity with the Woodcock-Muñoz (Woodcock, Muñoz-Sandoval, Ruef, & Alvarado, 2005) average score was .75 (Watson, 2004). We administered three passages at pretest and three alternative passages of the same reading level at posttest as a distal outcome.

DIBELS Nonsense Word Fluency (DIBELS NWF)

NWF is a standardized, individually administered test of pseudoword reading in English (Good & Kaminski, 2002). Similar to IDEL FPS, DIBELS NWF indicates knowledge of (a) letter-sound correspondences, in which letters represent their most common sounds; and (b) how to blend letter-sounds into whole units (i.e., pseudowords). According to Good and Kaminski (2002), alternate-form reliability coefficients for NWF ranged from .67 to .87. Studies have indicated moderate correlations (r = .56 – .65) between NWF at the end of kindergarten and the SAT-10 reading comprehension subtest at the end of first grade (Fien et al., 2008). We administered alternate forms of NWF at pretest and posttest as a distal outcome.

DIBELS Oral Reading Fluency 6th edition (DIBELS ORF)

ORF is a standardized, timed, individually administered test of accuracy and fluency with connected text (Good & Kaminski, 2002). Administration and scoring of the measure is the same as those of the IDEL FLO measure. In previous studies, alternate-form reliability coefficients of different reading passages from the same level of difficulty have ranged from .89 to .94 (Good & Kaminski, 2002). Correlations with the Texas Assessment of Knowledge and Skills (TAKS) taken in third grade range from .64 in Grade 1 to .68 in Grade 2 and .69 in Grade 3 (Wanzek et al., 2010). In this study, we administered three passages at pretest and three alternative passages of the same reading level at posttest as a distal outcome.

Student survey

This survey was adapted from a previous survey developed by Ward et al. (2011) and consisted of eight questions such as, *How much did you enjoy GraphoGame?* To reduce the possibility that students might have a difficult time reading the questions, a research assistant read the questions and asked students to color an icon that matched their thoughts. For example, a smiley face (☺) represented *A Lot*, a straight face (😐) represented *A Little*, and a sad face (☹) represented *None*. We used faces instead of words to reduce the possibility that some students would not understand exactly the meaning of the words either because of their low reading skills or because of their lack of vocabulary knowledge of basic constructs. For questions that asked for other answers (e.g., *When would you play GraphoGame*?), a picture of a house and a school were given. We administered the survey in Spanish to all the students in the treatment group (i.e., n = 48). A research assistant read each question aloud to the whole group and then waited until all students had responded to the question before reading the next one. In addition, 25% of the 48 students were randomly selected for an interview with open-ended questions such as, *How would you change the game to make it better?* We used this survey in combination with our observations to respond to RQ3.

Teacher survey

This adapted teacher survey from Ward et al. (2011) contained 29 questions; some used a Likert-type scale, while others were open-ended. Seven questions were specifically related to teacher perception of student engagement (e.g., *GraphoGame impacted students*. Potential ratings included *Negative, No Impact*, or *Positive*). As a follow-up, teachers were asked an open-ended question such as, *Please describe in detail the impact of the game for participating students*. Seventeen questions were related to the feasibility of implementing the game in an authentic classroom setting. Four additional questions taken from a national survey on educational technology focused on teacher perceptions of the use of technology in their classrooms. We used this survey to determine teacher satisfaction with GG-Spanish.

Quality of instruction

To respond to RQ4, reading instruction in treatment and comparison classrooms was observed. All teachers in control and treatment classrooms used the Spanish version, *Senderos*, of the Houghton Mifflin core reading program *Journeys* (Houghton Mifflin, 2010). Although the quality of instruction was not a factor in this analysis, we were interested in learning more about general classroom instruction for bilingual students. We used an instrument adapted from an instrument previously used in other projects that consisted of three parts: (a) teacher and school information; (b) eight items addressing the content of the instruction; and (c) a checklist of teacher behaviors that have been found to be effective instructional practices when teaching beginning reading in Spanish, such as providing an explanation of the task, modeling the activity, offering student opportunities to respond, and correcting errors (Coyne, Kame'enui, & Carnine, 2011). For each of the specified items in Part 2, behaviors were rated on a 4-point scale: *Consistently, Sometimes, Rarely*, or *Never*. No significant differences in the quality of Spanish instruction between treatment and control classrooms were observed.

Fidelity of implementation

Given that the intervention was a computer-based game and that student responses to the stimuli presented were recorded, results from our observations of student behaviors and the GG-Spanish data indicated that all students played the game with fidelity (i.e., they were fully engaged in the game throughout the intervention, and they listened to the sounds and responded promptly by clicking on stimuli presented visually onscreen). Our observations also indicated that all students adhered to the guidelines of the game during each session.

Data collection

Prior to pretesting, five data collectors received a half-day's training on the administration of the student measures by expert trainers. A one-hour refresher training was also provided prior to posttest data collection. After each training, a member of the research team and a data collector scored the same student assessment independently but in close proximity to each other. The percentage of agreement between two observers when allowing for 1-point discrepancy was 100% on all student measures. Student surveys regarding their level of engagement were conducted with the whole group. Interviews were conducted one-on-one, and student responses were written verbatim.

Data analysis procedure

To respond to RQ1 and RQ2, we first examined descriptively whether all students in the treatment group demonstrated greater improvements in their decoding and oral fluency skills compared to students in the control group in both Spanish and English from pretest to posttest. Then we examined descriptively whether risk status on our reading measures had a differential effect on outcomes (i.e., Did students who played GG-Spanish at a specific level of risk on the reading measures demonstrate greater improvement compared to students at other levels of risk? We also analyzed changes in gains from pretest (T1) to posttest (T2) within language (e.g., Spanish pretest to Spanish posttest) and *across* languages for evidence of cross-linguistic transfer with a mixed-model analysis (Murray, 1998) to account for the intraclass correlation associated with students nested within classrooms, the unit of assignment.

We fit models to the data with SAS PROC MIIXED version 9.2 (SAS Institute, 2009), using restricted maximum likelihood. Because the missing data were minimal and appeared to be at random (i.e., only two students did not complete one or more of the pretest or posttest assessments), we used list-wise deletion for analysis of our proximal and distal outcome measures in Spanish and in English. To ease interpretation, we computed an effect size, Hedges's g (Hedges, 1981), for each effect according to What Works Clearinghouse (WWC, 2014) standards. Hedges's g is an individual-level effect size comparable to Cohen's d (Cohen, 1988).

Results

Evaluation of GraphoGame

We present descriptive statistics for students at varying levels of risk at pretest and posttest in Tables 2 and 3 respectively to illustrate the variability in scores among groups at both time points. Findings suggest that at pretest, as illustrated in Table 2, differences between the treatment and control group were minimal on Spanish pseudoword reading (IDEL FPS), Spanish Oral Reading Fluency (IDEL FLO), and English pseudoword reading (DIBELS NWF), but they were substantial in English ORF (DIBELS ORF), favoring the control group (M = 26.57 vs. 37.91 words per minute respectively). Differences in scores between the treatment and control group were also similar at posttest. Students in the treatment group and control group performed equally well on IDEL FPS, IDEL FLO, and DIBELS NWF, but differences remained between the treatment and control group on DIBELS ORF (M = 39.57 vs. 49.73 words per minute respectively), favoring the control group.

Our descriptive analysis also revealed the large variability of reading skills in Spanish and in English between students across the treatment and control groups at pretest and posttest. For example, as illustrated in Table 2, the range of scores on IDEL FLO at pretest was between 0 and 129 and between

Table 2. Descriptive statistics by group for pretest outcomes (T1) in Spanish and English.

	Treatment				Control			
	N	M (SD)	Min	Max	N	M (SD)	Min	Max
Spanish								
IDEL FPS	46	89.33 (36.93)	24	170	32	92.72 (50.55)	7	207
At Risk	4	31.00 (7.17)	24	39	5	23.20 (11.19)	7	35
Some Risk	12	54.75 (6.31)	46	63	7	52.14 (9.41)	41	68
Low Risk	30	110.93 (25.47)	72	170	20	124.30 (34.44)	73	207
IDEL FLO	46	35.35 (17.56)	2	94	32	41.13 (26.80)	0	129
At Risk	4	3.25 (1.26)	2	5	5	4.60 (4.22)	0	10
Some Risk	12	26.83 (11.07)	15	43	7	45.17 (25.40)	15	84
Low Risk	30	43.03 (14.09)	20	94	20	50.10 (21.84)	26	129
English								
DIBELS NWF	46	54.11 (30.66)	11	117	32	48.72 (31.41)	0	132
At Risk	12	20.08 (5.60)	11	27	10	16.80 (11.55)	0	29
Some Risk	13	40.85 (5.52)	32	49	8	38.38 (6.50)	30	49
Low Risk	21	81.76 (21.60)	50	117	14	77.43 (22.36)	51	132
DIBELS ORF	48	26.57 (21.00)	0	71	32	37.91 (29.38)	0	103
At Risk	12	7.08 (7.50)	0	27	10	23.10 (29.53)	0	93
Some Risk	13	31.85 (20.69)	5	69	8	36.75 (25.47)	5	76
Low Risk	21	34.43 (19.85)	6	71	14	49.14 (28.28)	8	103

Note. IDEL FPS = Fluidez en las Palabras sin Sentido; FLO = Fluidez en la Lectura Oral; DIBELS NWF = Nonsense Word Fluency; DIBELS ORF = Oral Reading Fluency.

Risk status is based on IDEL FPS for Spanish pretest outcomes (At Risk = 0–39, Some Risk = 40–69, Low Risk = 70 or greater) and DIBELS NWF for English pretest outcomes (At Risk = 0–29, Some Risk = 30–49, Low Risk = 50 or greater) in the middle of first grade.

Table 3. Descriptive statistics by group for posttest outcomes (T2) in Spanish and English.

	Treatment				Comparison			
	N	M (SD)	Min	Max	N	M (SD)	Min	Max
Spanish								
IDEL FPS	46	118.80 (42.88)	34	203	32	123.47 (53.93)	4	205
At Risk	8	52.25 (12.09)	34	69	5	40.80 (25.09)	4	67
Some Risk	4	81.25 (6.85)	73	89	4	79.50 (3.70)	75	84
Low Risk	34	138.88 (28.50)	90	203	23	149.09 (37.30)	91	205
IDEL FLO	46	47.76 (19.68)	8	93	32	50.78 (25.45)	0	107
At Risk	8	21.00 (10.77)	8	41	5	18.40 (19.01)	0	46
Some Risk	4	38.25 (17.25)	25	62	4	44.75 (38.13)	15	100
Low Risk	34	55.18 (15.34)	26	93	23	58.87 (18.32)	32	107
English								
DIBELS NWF	46	71.89 (33.58)	16	135	32	72.25 (40.52)	0	138
At Risk	6	19.67 (2.88)	16	22	4	17.50 (14.15)	0	29
Some Risk	7	38.71 (6.32)	30	47	10	39.70 (5.79)	32	47
Low Risk	33	88.42 (23.30)	50	135	18	102.50 (25.35)	53	138
DIBELS ORF	46	39.57 (24.47)	0	88	32	49.73 (30.84)	0	137
At Risk	6	7.83 (5.98)	0	16	4	25.50 (28.24)	0	56
Some Risk	7	19.71 (11.10)	7	38	10	33.90 (25.90)	5	97
Low Risk	33	49.52 (20.89)	13	88	18	63.92 (27.19)	22	137

Note. IDEL FPS = Fluidez en las Palabras sin Sentido; FLO = Fluidez en la Lectura Oral; DIBELS NWF = Nonsense Word Fluency; DIBELS ORF = Oral Reading Fluency.

Risk status is based on IDEL FPS for Spanish posttest outcomes (At Risk = 0–69, Some Risk = 70–89, Low Risk = 90 or greater) and DIBELS NWF for English posttest outcomes (At Risk = 0–29, Some Risk = 30–49, Low Risk = 50 or greater) at the end of first grade.

0 and 103 on DIBELS ORF without taking group assignment and risk status into account. Similarly, as illustrated in Table 3, at posttest scores on oral reading fluency ranged from 0 to 107 in Spanish and from 0 to 137 in English without taking group assignment and risk status into account.

Descriptive analysis based on risk status

When examining student outcomes by risk status, findings indicated that the percentage of students at low risk on Spanish pseudoword reading increased from 65% at pretest to 74% at posttest. Similar trends were also observed by students in the control group in which 63% of students were at low risk at pretest and 72% were at low risk at posttest. In English, however, a larger percentage of students in the treatment group moved on DIBELS NWF to the low-risk category from pretest to posttest compared to the control group (from 46% to 72% in the treatment group vs. 44% to 56% in the control group). Results on DIBELS ORF did not differ between groups.

Moreover, as illustrated in Table 4, it appears that students who were identified as being at some risk on IDEL FPS at pretest were those who benefited the most in English pseudoword reading from playing GG-Spanish. For example, of the 12 students in the treatment group who

Table 4. Descriptive analyses of student movement across risk levels (Spanish pseudoword reading at pretest and English pseudoword reading at posttest).

	DIBELS NWF Posttest Risk Status					
	Treatment			Control		
IDEL FPS Pretest Risk Status (n)	At Risk	Some Risk	Low Risk	At Risk	Some Risk	Low Risk
At Risk (*nt* = 4; *nc* = 5)	4 (100%)	0 (0.0%)	0 (0.0%)	2 (40.0%)	2 (40.0%)	1 (20.0%)
Some Risk (*nt* = 12; *nc* = 7)	2 (16.7%)	3 (25.0%)	7 (58.3%)	0 (0.0%)	6 (85.7%)	1 (14.3%)
Low Risk (*nt* = 30; *nc* = 20)	0 (0.0%)	4 (13.3%)	26 (86.7%)	2 (10.0%)	2 (10.0%)	16 (80.0%)

Note. *n* = the total number of students in the treatment (t) and control groups (c) identified as at risk, some risk, and low risk at pretest, (i.e., in the middle of first grade) on the Spanish pseudoword measure (IDEL FPS; At Risk = 0–39, Some Risk = 40–69, Low Risk = 70 or greater); DIBELS NWF Posttest Risk Status refers to students who were at risk, some risk, or low risk on English pseudoword reading at the end of first grade (At Risk = 0–29, Some Risk = 30–49, Low Risk = 50 or greater).

were at some risk on IDEL FPS at pretest, seven (58%) were later at low risk on DIBELS NWF, compared to only one student in the control group (14.3%).

Analysis of gains

After our initial descriptive analysis, we conducted a statistical analysis of gains. As illustrated in the first two columns of Table 5, examination of differences between conditions in gains on IDEL FPS and FLO from pretest to posttest revealed no statistically significant differences between the treatment and comparison groups ($p = 0.81$ and $p = 0.14$ respectively). However, the Hedges's g value of 0.30 for FLO gains indicate a small but potentially important effect.

The next two columns in Table 5 present results on the English measures from pretest to posttest. Hedges's g values for gains on DIBELS NWF and ORF are small, and the negative effect for DIBELS NWF indicates that students in the treatment group had lower scores at pretest, and their gains on NWF did not surpass the gains made by students in the control group. The last two columns in Table 5 present results of cross-linguistic transfer from IDEL FPS at pretest to DIBELS NWF at posttest (i.e., cross-linguistic transfer of pseudoword reading) and from IDEL FLO at pretest to DIBELS ORF at posttest (i.e, cross-linguistic transfer of oral reading fluency skills). Results indicate that condition was a small positive, but not significant, predictor of the gains students made on English pseudoword reading ($p = .73$, $g = 0.15$). On the other hand, condition appeared to be a small negative, but not significant, predictor of the gains students made on English oral reading fluency ($p = .46$, $g = -0.31$).

Table 5. Fixed effects, variance components, and specific effect estimates for main effects of condition on reading gains.

		Spanish Measures		English Measures		Cross-Linguistic Transfer	
Effect or Statistic		FPS	FLO	NWF	ORF	NWF	ORF
Fixed effects	Classroom	30.81** (4.08)	9.86** (1.00)	23.58* (6.33)	11.87* (2.04)	−21.52* (8.79)	8.59 (3.79)
	Condition	−1.39 (5.34)	2.62 (1.33)	−5.60 (8.21)	1.10 (2.65)	4.29 (11.37)	−4.19 (4.90)
Variance components	Classroom	−14.28 (28.99)	−2.81 (2.08)	33.88 (68.33)	0.93 (7.00)	106.78 (126.73)	17.85 (24.23)
	Residual	766.05*** (126.64)	78.61*** (13.02)	734.77*** (121.72)	118.24*** (19.56)	753.13*** (124.64)	171.18*** (28.36)
Hedges' *g*	Condition	−0.05	0.30	−0.20	0.10	0.15	−0.31
p value	Condition	0.81	0.14	0.54	0.71	0.73	0.46
Improvement index	Condition	−2.01%	11.80%	−8.08%	4.03%	5.96%	−12.20%
ICC			<.0001	.044	.008	0.12	0.09

Note. DIBELS NWF = Nonsense Word Fluency, DIBELS ORF = Oral Reading Fluency; IDEL FPS = Fluidez en las Palabras sin Sentido; IDEL FLO = Fluidez en la Lectura Oral. Covariates in columns 3 and 4 were FPS and FLO at pretest, in columns 5 and 6 were NWF and ORF at pretest. In columns 7 and 8, covariates were FPS at pretest for NWF outcomes, and FLO at pretest for ORF outcomes. ***$p < .001$, **$p < .01$, *$p < .0$.

Student engagement with GraphoGame

In general, more than 94% of students believed that GG-Spanish helped them with reading words, and 96% indicated that they enjoyed playing GraphoGame a lot. In addition, more than 60% of students felt that they played GraphoGame the right amount of time, and 100% of students indicated that they were more excited about reading after playing GraphoGame. During the interview, 58% (7/12) said they would not change anything in the game; one student wanted to *learn more*, one student wanted to *make it longer*, one student indicated that s/he wanted to play the game in English, and one student wanted *more games*. Results of the survey illustrated in Table 6 aligned with teacher beliefs about student use of the game.

Teacher perception of student level of engagement

Results of the teacher survey indicated that teachers thought students enjoyed playing the game and were more excited about reading after playing. One of the teachers said, "My students were so excited to log in and play GraphoGame. Over time, I even saw that my low-level students were excited to participate and identify letter names and sounds." In addition, teachers unanimously acknowledged that GraphoGame served as a significant tool in different forms in their language arts class and found it easy to incorporate as a means for providing students with opportunities to independently practice their letter-sound correspondence knowledge while teachers provided small-group instruction to other students. Finally, all teachers expressed their willingness to continue using GraphoGame in their classrooms and that they would recommend the game to other teachers.

Discussion

The purpose of this study was to conduct a small-scale evaluation of a CARI to explore the cross-linguistic transfer of early reading skills from Spanish to English. We hypothesized that if a CARI could potentially increase the decoding skills of bilingual students in Spanish and English, and consequently their oral reading fluency, teachers could devote more of their instructional time during whole-group and small-group instruction to teaching vocabulary and comprehension in either Spanish or English (i.e., depending on the bilingual program) instead of a discrete skill such as decoding. In addition, the CARI could free some time for teachers to work with struggling bilingual students in small groups in a Response-to-Intervention approach, and it could provide all bilingual students, including those who were not identified as struggling, additional practice opportunities to strengthen their alphabetic understanding and decoding skills in an engaging instructional environment.

Table 6. Student engagement survey results by language of instruction (n = 46).

Survey Question	☹	😐	☺
How much did GraphoGame help you with reading words?	0 (0%)	3 (6.5%)	43 (93.5%)
How much did you enjoy GraphoGame?	0 (0%)	2 (4.3%)	44 (95.7%)
The amount of time I played GraphoGame was …	2 (4.3%)	29 (63.0%)	15 (32.6%)
How well could you hear the sounds and letters of the words?	0 (0%)	13 (28.3%)	33 (71.7%)
Clicking on the words and letters was …	1 (2.2%)	3 (6.5%)	42 (91.3%)
Now that I have played GraphoGame …	0 (0%)	0 (0%)	46 (100%)
Did you understand what you were supposed to do in each of the GraphoGame activities?	0 (0%)	1 (2.2%)	45 (56.3%)

Note. For each item, ☹ represented the most negative response on the scale, while ☺ represented the most positive response.

Our results were mixed. We found some potentially meaningful educational effects for English pseudoword reading when taking Spanish pseudoword reading at pretest into account. In contrast to the positive pseudoword reading effect in English, we found a negative effect of GG-Spanish on English oral reading fluency at posttest when taking Spanish oral reading fluency at pretest into account. Next, we discuss our findings in the context of using computer games to improve learning.

Conducting evaluations using an experimental design

One plausible explanation for observing nonsignificant effects of GG-Spanish on student outcomes could have been our small sample size given that we used the classroom as the unit of analysis. Our decision to assign classrooms to conditions instead of students within classrooms was to remove the complexity of teachers having to spend time training a subset of students on what to do while their peers were playing the game and to mitigate students feeling that they had been treated unfairly by not letting them play the game as did their peers.

Another influential factor in our lack of effects might have been the characteristics of teachers in the control group. From our observations, teachers in the control group spent a substantial amount of time teaching decoding skills, and they may have been highly effective based on their student outcomes. Moreover, all teachers understood the purpose of the study, and thus the control teachers might have worked more than normal on building their student decoding and oral reading fluency skills, thereby diluting the effects of GG-Spanish.

Closer examination of the instructional design of GG-Spanish

Another factor that may have influenced the results concerns the relation between GG–Spanish and Spanish orthography. A close examination of the content of GG-Spanish order of presentation of sound and syllable patterns revealed that these patterns did not always reflect the characteristics of the Spanish orthographic system. For example, some of the syllables presented in the game do not exist in Spanish (e.g., *bier, bres, trein, triun, grien, pli, nus)*, or they are very low-frequency syllables that appear in less-frequent words (e.g., *hia*, as in *hiato* ["hiatus" in English], or *tua* as in *tatuaje* ["tattoo" in English]). Consequently, the game might not have provided students with the necessary scaffolds to identify these syllables automatically. For instance, according to research conducted in Argentina and Spain (see Cuetos, 2010; Signorini, 1997), the most frequent syllable patterns in Spanish are CV, CVC, V, and CVV. GG-Spanish however, used CV, VC, CVC, and CVV patterns (i.e., it included VC patterns before CVC, and it did not include the one-vowel syllable pattern before CVV). Moreover, some of the CVV syllables included the silent *u* (e.g., *qui, gue*), which is confusing for students in first grade unless they have been explicitly taught the letter-sound correspondence rule for the silent *u*. Thus, a potential influential variable that might have reduced the efficacy of GG-Spanish could be related to the design of the game. This hypothesis, however, would need to be tested further with different iterations of the presentation of sound and syllable patterns.

Individual reading characteristics of bilingual students

This study also shed light on the large differences in beginning reading skills in Spanish and in English of students attending bilingual schools in a large urban high-poverty neighborhood. As illustrated in Table 3, some students were reading 107 words per minute in Spanish and 137 words per minute in English at the end of first grade, while other students could not read any words in neither English nor Spanish. Furthermore, our results suggest that approximately 35% of students in the sample were struggling in decoding in Spanish at the end of first grade and, to our knowledge, had not been provided with the supports they needed to improve their beginning reading skills. Although reading words fast is not the main goal of teaching oral reading fluency in the lower grades, reading words automatically in isolation and fluently in connected text has a large significant

correlation with reading comprehension in Spanish and in English (Baker et al., 2011; Crespo, 2014). Thus, although it is likely that CARIs that incorporate evidence-based principles of instructional design to scaffold student learning could support student acquisition of beginning reading skills (see, for example, the Cheung & Slavin, 2013 synthesis and the experimental study on the Finnish version of Graphogame by Saine et al., 2011), additional empirical research is needed to determine the effectiveness of these games in other languages, before they are used across schools.

Cross-linguistic transfer

Regarding cross-linguistic transfer, our findings suggest that some cross-linguistic transfer might have taken place as a consequence of students using GG-Spanish, but these findings need to be corroborated further with a larger study. Using a CARI to examine more closely cross-linguistic transfer is ideal because it allows researchers to isolate the language of the game from other potential variables, such as student behavior, teacher years of experience, and teacher proficiency with English and Spanish (variables that are likely to influence outcomes in the context of teacher-delivered small-group instruction). Experimental studies such as the one we conducted with a larger sample size could also advance this field of research, particularly given that most studies on cross-linguistic transfer are correlational (Melby-Lervåg & Lervåg, 2011).

Usability and engagement of GraphoGame

Our study also confirmed previous findings indicating that a CARI such as GraphoGame is highly engaging and appealing to students. Moreover, the game required minimal training and was easy for students to understand and follow. Children were able to use the mouse with ease and independently kept track of the number of minutes their peers were playing the game without teacher support, and they were always aware of whose turn it was to play the game when it was time to change activities. Only a few students out of the 78 (less than 10%) wanted the game to be funnier or more challenging.

Feasibility of GraphoGame

All teachers thought the game would work well in their classrooms, and the processes required to download it were reasonable. Teachers also commented that in a Response-to-Intervention approach, the use of GraphoGame as an independent activity allowed them to spend more time with their struggling students without interruptions from other students in the class because students were able to keep track of the amount of time each of them had to play, and they were engaged for the full 10 minutes that they played GG-Spanish.

Conclusions

This experimental study is the first one conducted in the United States using a CARI with bilingual students in a bilingual setting taking their reading skills in both languages into account. The study is a low-cost, low-stakes experimental study that was intended to obtain an initial understanding of whether a computer game to practice decoding could build the reading skills of bilingual students in their native language, Spanish, and in their second language, English. Our results were mixed, and although some cross-linguistic transfer from Spanish to English might have occurred, particularly in pseudoword reading, more research needs to be conducted to confirm these findings.

Computer games are being developed at a fast pace to support student literacy and language acquisition, and GraphoGame in particular is currently being used by thousands of children across the globe to build their early literacy skills. In our study, teachers indicated that it could fit very well in a Response-to-Intervention approach to support all students in the classroom. However, the benefits of computer games, such as GG-Spanish, is not yet clear. Thus, given the attractiveness of

the use of computer games in schools, seeking funding to conduct future iterative design trials to improve the content of computer games or CARIS and carrying out experimental studies to determine the effects of these games on student learning outcomes warrants further investigation.

Acknowledgments

We would like to acknowledge the contributions to this project by Shama Rajani.

Funding

The research reported here was supported, in part, by grants from the National Science Foundation (DRL 1254189 to Eric Hamilton at Pepperdine University and DRL 0733323 to Ron Cole and Wayne Ward at Boulder Learning Inc.). The opinions expressed are those of the authors and do not represent the views of the National Science Foundation.

References

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: Massachusetts Institute of Technology.

Bae, J. (2007). Development of English skills need not suffer as a result of immersion: Grades 1 and 2 writing assessment in a Korean/English two-way immersion program. *Language Learning*, *57*, 299–332. doi:10.1111/lang.2007.57.issue-2

Baker, D. L., Basaraba, D., & Polanco, P. (2016). Connecting the present to the past: Furthering the research on bilingualism and bilingual education. *Review of Research in Education*, *40*(1), 821–883. doi:10.3102/0091732X16660691

Baker, D. L., Cummings, K. D., Good, R. H., & Smolkowski, K. (2007). *Indicadores Dinámicos del Éxito in la Lectura (IDEL®): Summary of decision rules for intensive, strategic, and benchmark instructional recommendations in kindergarten through third grade*. Technical Report No. 1. Eugene, OR: Dynamic Measurement Group. Retrieved from https://dibels.uoregon.edu/docs/benchmarkgoals_idel_3x.pdf

Baker, D. L., Good, R. H., Knutson, N., & Watson, J. M. (Eds.). (2006). *Indicadores Dinámicos del Éxito en la Lectura* (7th ed.). Eugene, OR: Dynamic Measurement Group.

Baker, D. L., Good, R. H., Mross, A. P., McQuilkin, E., Watson, J., Chaparro, E. A., … Sanford, A. K. (2006). Fluidez en la Lectura Oral IDEL. In D. L. Baker, R. H. Good, N. Knutson, & J. M. Watson (Eds.), *Indicadores Dinámicos del Éxito en la Lectura* (7th ed.). Eugene, OR: Dynamic Measurement Group. Retrieved from http://dibels.uoregon.edu/

Baker, D. L., Good, R. H., Peyton, J., & Watson, J. (2004). *Alternate form reliability of IDEL Fluidez en las Palabras sin Sentido*. Raw data. Eugene, OR: University of Oregon.

Baker, D. L., Park, Y., & Baker, S. K. (2010). Effects of initial status and growth in pseudoword reading on Spanish reading comprehension at the end of first grade. *Psicothema*, *22*, 955–962.

Baker, D. L., Park, Y., & Baker, S. K. (2012). The reading performance of English learners in grades 1–3: The role of initial status and growth on reading fluency in Spanish and English. *Reading & Writing: An Interdisciplinary Journal*, *25*, 251–281. doi:10.1007/s11145-010-9261-z

Baker, D. L., Park, Y., & Baker, S. K. (2013, July). *Effect of English language proficiency and Spanish and English literacy on English reading comprehension for Spanish-speaking English learners*. Interactive paper presented at the Society for the Scientific Study of Reading Conference, Hong Kong, China.

Baker, D. L., Stoolmiller, M., Good, R. H., & Baker, S. K. (2011). Effects of reading comprehension on passage fluency in Spanish and English for second-language learners. *School Psychology Review*, *40*(3), 331–351.

Cheung, A. K., & Slavin, R. E. (2013). Effects of educational technology applications on reading outcomes for struggling readers: A best-evidence synthesis. *Reading Research Quarterly*, *48*, 277–299. doi:10.1002/rrq.50

Clemens, N. H., Shapiro, E. S., Wu, J., Taylor, A. B., & Caskie, G. L. (2014). Monitoring early first-grade reading progress: A comparison of two measures. *Journal of Learning Disabilities*, *47*, 254–270. doi:10.1177/0022194124544455

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Coyne, M. D., Kame'enui, E. J., & Carnine, D. W. (Eds.). (2011). *Effective teaching strategies that accommodate diverse learners* (4th ed.). Upper Saddle River, NJ: Pearson.

Crespo, P. (2014). *Modelo de respuesta a la intervención: Nivel 2 de prevención de las dificultades de aprendizaje en lectura en la comunidad autónoma de Canarias* (Unpublished doctoral dissertation). Universidad de la Laguna, Tenerife, Islas Canarias.

Cuetos, F., (2010). *Psicología de la lectura*. España: Wolters Kluwer, S.A.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, *49*, 222–25l. doi:10.2307/1169960

Davies, R., Cuetos, F., & Glez-Seijas, R. (2007). Reading development and dyslexia in a transparent orthography: A survey of Spanish children. *Annals of Dyslexia*, *57*, 179–198. doi:10.1007/s11881-007-0010-1

Ehri, L. (2005). Learning to read word: Theory, findings, and issues. *Scientific Studies of Reading*, *9*, 167–188. doi:10.1207/s1532799xssr0902_4

Fien, H., Baker, S. K., Smolkowski, S., Mercier Smith, J. L., Kame'enui, E. J., & Beck, C. T. (2008). Using Nonsense Word Fluency to predict reading proficiency in kindergarten through second grade for English learners and native English speakers. *School Psychology Review*, *37*, 391–408.

Fien, H., Park, Y., Baker, S. K., Smith, J. L., Stoolmiller, M., & Kame'enui, E. J. (2010). An examination of the relation of Nonsense Word Fluency initial skill status and growth to reading outcomes for beginning readers. *School Psychology Review*, *39*, 631–653.

Foorman, B. R., Breier, J. L., & Fletcher, J. M. (2003). Interventions aimed at improving reading success: An evidence-based approach. *Developmental Neuropsychology*, *24*, 613–639. doi:10.1207/S15326942DN242&3_06

Frost, R. (2005). Orthographic systems and skilled word recognition processes in reading. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 272–295). Malden, MA: Blackwell.

Genessee, F., Geva, E., Dressler, C., & Kamil, M. (2006). Synthesis: Cross-linguistic relationships. In D. August & T. Shanahan (Eds.), *Developing literacy in second language learners: Report of the National Reading Panel on Language-Minority Children and Youth* (pp. 153–174). Mahwah, NJ: Lawrence Erlbaum.

Gerber, M., Jimenez, T., Leafstedt, J., Villaruz, J., Richards, C. & English, J. (2004). English reading effects of small-group intensive intervention in Spanish for K-1 English learners. *Learning Disabilities Research & Practice*, *19*(4), 239–251.

Good, R. H., & Kaminski, R. A. (2002). *Dynamic indicators of basic early literacy skills* (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education (RASE)*, *7*, 6–10. doi:10.1177/074193258600700104

Hall, T. E., Hughes, C. A., & Filbert, M. (2000). Computer assisted instruction in reading for students with learning disabilities: A research synthesis. *Education and Treatment of Children*, *23*, 173–193.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. J*ournal of Educational Statistics*, *6*, 107–128. doi: 10.3102/10769986006002107

Houghton Mifflin. (2010). *Senderos*. ISBN-10: 0547243421.

Jere-Folotiya, J., Chansa-Kabali, T., Munachaka, J. C., Sampa, F., Yalukanda, C., Westerholm, J., & Lyytinen, H. (2014). The effect of using a mobile literacy game to improve literacy levels of grade one students in Zambian schools. *Educational Technology Research & Development*, *62*, 417–436. doi:10.1007/s11423-014-9342-9

Kame'enui, E. J., & Simmons, D. C. (2001). Introduction to this special issue: The DNA of reading fluency. *Scientific Studies of Reading*, *5*, 203–210. doi:10.1207/S1532799XSSR0503_1

Kamykowska, J., Haiman, E., Latvala, J.-M., Richardson, U., & Lyytinen, H. (2013). Developmental changes of early reading skills in six-year-old Polish children and GraphoGame as a computer-based intervention to support them. *L1—Educational Studies in Language and Literature, 13*, 1–13. doi:10.17299/L1ESLL-2013.01.05

Kyle, F., Kujala, J., Richardson, U., Lyytinen, H., & Goswami, U. (2013). Assessing the effectiveness of two theoretically motivated computer-assisted reading interventions in the United Kingdom: GG Rime and GG Phoneme. *Reading Research Quarterly*, *48*, 61–76. doi:10.1002/rrq038

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*, 293–323. doi:10.1016/0010-0285(74)90015-2

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor, MI: University of Michigan Press.

Lan, Y., Sung, Y., & Chang, K. (2009). Let us read together: Development and evaluation of a computer-assisted reciprocal early English reading system. *Computers & Education*, *53*, 1188–1198. doi:10.1016/j.compedu.2009.06.002

Luo, Y. C., Chen, X., & Geva, E. (2014). Concurrent and longitudinal cross-linguistic transfer of phonological awareness and morphological awareness in Chinese-English bilingual children. *Written Language & Literacy*, *17* (1), 89–115. doi:10.1075/wll.17.1.05luo

Lyon, G. R. (2009). *The NICHD research program in reading development, reading disorders, and reading instruction*. Washington, DC: National Center for Learning Disabilities. Retrieved from http://www.ncld.org/ld-basics/ld-aamp-lang/reading

Lyytinen, H., Ronimus, M., Alanko, A., Poikkeus, A., & Taanila, M. (2007). Early identification of dyslexia and the use of computer-based game practice to support reading acquisition. *Nordic Psychology*, *59*, 109–126. doi:10.1027/1901-2276.59.2.109

Macaruso, P., & Rodman, A. (2011). Benefits of computer-assisted instruction to support reading acquisition in English language learners. *Bilingual Research Journal*, *34*, 301–315. doi:10.1080/15235882.2011.622829

Maxwell, L. (2012, March). Dual classes see growth in popularity. *Education Week*, *31*(26).

Melby-Lervåg, M., & Lervåg, A. (2011). Cross-linguistic transfer of oral language, decoding, phonological awareness, and reading comprehension: A meta-analysis of the correlational evidence. *Journal of Research in Reading*, *34*, 114–135. doi:10.1111/j.1467-9817.2010.01477/x

Murray, D. M. (1998). *Design and analysis of group-randomized trials*. New York: Oxford University Press.

Nakamoto, J., Lindsey, K. A., & Manis, F. R. (2012). Development of reading skills from K–3 in Spanish-speaking English language learners following three programs of instruction. *Reading and Writing: An Interdisciplinary Journal*, *25*, 537–567. doi:10.1007/s11145-010-9285-4

Nemeth, K., & Simon, F. (2013). Using technology as a teaching tool for dual language learners in preschool through grade 3. *Young Children*, *68*(1), 48–52.

Ojanen, E., Ronimus, M., Ahonen, T., Chansa-Kaball, T., February, P., Jere-Folotiya, J., …Lyytinen, H. (2015). GraphoGame - a catalyst for multi-level promotion of literacy in diverse contexts. *Frontiers in Psychology*, *6*, 671. doi: 10.3389/fpsyg.2015.00671.

Perfetti, C. A. (1985). *Reading ability*. New York, NY: Oxford University.

Plasencia-Peinado, J., Baker, D. L., Good, R. H., & Peinado, R. (2006). Fluidez en las Palabras sin Sentido. In D. L. Baker, R. H. Good, N. Knutson, & J. M. Watson (Eds.), *Indicadores Dinámicos del Éxito en la Lectura* (7th ed.). Eugene, OR: Dynamic Measurement Group. Retrieved from http://dibels.uoregon.edu/

Potocki, A., Ecalle, J., & Magnan, A. (2013). Effects of computer-assisted comprehension training in less-skilled comprehenders in second grade: A one-year follow-up study. *Computers & Education*, *63*, 131–140. doi:10.1016/j.compedu.2012.12.011

Ronimus, M., & Richardson, U. (2014). Digital game-based training of early reading skills: Overview of the GraphoGame method in a highly transparent orthography/Entrenamiento de habilidades de lectura temprana basado en un juego digital: Visión general del método GraphoGame en una ortografía transparente. *Estudios De Psicología*, *35*, 648–661. doi:10.1080/02109395.2014.974424

Rusu-Bodea, A. S. (2016, July). E-learning tools' influence over students' motivation for English as a second language acquisition. In *The international scientific conference elearning and software for education* (Vol. 3, p. 131). "Carol I" National Defence University

Saine, N. L., Lerkkanen, M., Ahonen, T., Tolvanen, A., & Lyytinen, H. (2011). Computer-assisted remedial reading intervention for school beginners at risk for reading disability. *Child Development*, *82*, 1013–1028. doi:10.1111/j.1467-8624.2011.01580.x

San Francisco, A. R., Carlo, M. S., August, D., & Snow, C. E. (2006). The role of instruction and vocabulary in the English phonological awareness of Spanish-English bilingual children. *Applied Psycholinguistics*, *27*, 229–246. doi:10.1017/S0142716406060267

SAS Institute Inc. (2009). *SAS/STAT ®9.2 user's guide*. Cary, NC: SAS Institute Inc.

Seymour, P. H. K. (2005). Early reading development in European orthographies. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 296–315). Malden, MA: Blackwell.

Signorini, A. (1997). Word reading in Spanish: A comparison between skilled and less skilled beginning readers. *Applied Psycholinguistics*, *18*, 319–344.

Stanovich, K. E. (1986). Matthew Effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*, 360–407. doi:10.1598/RRQ.21.4.1

State Educational Technology Directors Association (SETDA). (2010). *National educational technology trends: 2010*. Retrieved from www.setda.org

University of Oregon Center on Teaching & Learning. (2012). *DIBELS 6th edition benchmark goals: Three assessment periods per year*. Retrieved from https://dibels.uoregon.edu/docs/benchmarkgoals.pdf

Verhoeven, L. T. (2011). Second language reading acquisition. In M. Kamil, D. Pearson, E. Moje, & P. Afflerbach (Eds.), *Handbook of reading research IV* (pp. 661–683). New York, NY: Routledge.

Wanzek, J., Roberts, G., Linan-Thompson, S., Vaughn, S., Woodruff, A., & Murray, C. (2010). Differences in the relationship of oral reading fluency and high-stakes measures of reading comprehension. *Assessment for Effective Intervention*, *35*(2), 67–77. doi:10.1177/1534508409339917

Ward, W., Cole, R., Bolanos, D., Buchenroth-Martin, C., Svirsky, E., Vuuren, S. V., … Becker, L. (2011). My Science Tutor: A conversational multimedia virtual tutor for elementary school science. *ACM—Transactions on Speech and Language Processing*, *7*, 1–29. doi:10.1145/1998384.1998392

Watson, J. M. (2004). *Examining the reliability and validity of the Indicators Dinámicos del Éxito en la Lectura: A research study* (Unpublished doctoral dissertation). University of Oregon, Eugene.

What Works Clearinghouse. (2014). *Procedures and standards handbook* (Version 3.0). Washington, DC: U.S. Department of Education, Institute of Education Sciences. Retrieved from the Institute of Education Sciences, National Center for Education Evaluation, WWC: http://ies.ed.gov/ncee/wwc/

Williams, K. T., Cassidy, J., & Samuels, S. J. (2001). *Group assessment and diagnostic evaluation*. New York, NY: Pearson.

Woodcock, R., Muñoz-Sandoval, A., Ruef, M., & Alvarado, C. (2005). *Woodcock-Muñoz language survey*. Riverside Publishing Company.